

#25

RECEIVED

MAR 13 2003

TECH CENTER 1600/2900



ENTERED

1600

RAW SEQUENCE LISTING

DATE: 03/06/2003 P6

PATENT APPLICATION: US/09/724,000A

TIME: 12:42:46

Input Set : A:\00450Aseq.txt

Output Set: N:\CRF4\03062003\I724000A.raw

```

3 <110> APPLICANT: Polverino, Anthony J.
4   Luethy, Roland
6 <120> TITLE OF INVENTION: Secreted Epithelial Colon Stromal-1 Molecules and Uses
7   Thereof
9 <130> FILE REFERENCE: 00-450-A
11 <140> CURRENT APPLICATION NUMBER: 09/724,000A
12 <141> CURRENT FILING DATE: 2000-11-28
14 <150> PRIOR APPLICATION NUMBER: 09/599,087
15 <151> PRIOR FILING DATE: 2000-06-21
17 <160> NUMBER OF SEQ ID NOS: 22
19 <170> SOFTWARE: PatentIn Ver. 2.0
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 744
23 <212> TYPE: DNA
24 <213> ORGANISM: Mus musculus
26 <220> FEATURE:
27 <221> NAME/KEY: CDS
28 <222> LOCATION: (38)..(274)
30 <220> FEATURE:
31 <221> NAME/KEY: sig_peptide
32 <222> LOCATION: (38)..(109)
34 <400> SEQUENCE: 1
35 gcttcctccc taggcgtgag actcgggctc cttcact atg aga ctt cta gcc ctt 55
36                                     Met Arg Leu Leu Ala Leu
37                                     1           5
39 tcc ggt ctg ctc tgc atg ctg ctc ctc tgt ttc tgc att ttc tcc tca 103
40 Ser Gly Leu Leu Cys Met Leu Leu Leu Cys Phe Cys Ile Phe Ser Ser
41           10           15           20
43 gaa ggg aga aga cat cct gcc aag tcc ttg aaa ctc agg cgc tgc tgt 151
44 Glu Gly Arg Arg His Pro Ala Lys Ser Leu Lys Leu Arg Arg Cys Cys
45           25           30           35
47 cac cta tct cct aga tcc aag ctg aca acc tgg aaa gga aac cac aca 199
48 His Leu Ser Pro Arg Ser Lys Leu Thr Thr Trp Lys Gly Asn His Thr
49           40           45           50
51 agg ccc tgc aga ctc tgc aga aac aag cta cca gtc aag tca tgg gtg 247
52 Arg Pro Cys Arg Leu Cys Arg Asn Lys Leu Pro Val Lys Ser Trp Val
53 55           60           65           70
55 gtg cct ggg gct ctc cca cag ata tag ggcctcctcc gccagatga 294
56 Val Pro Gly Ala Leu Pro Gln Ile
57           75
59 agcgttgatg ccagatgtg gagacaccag aagcatacac actatgttgc cttgccctt 354
61 gccaatgagc tgtgacactg gaatgcttca cttcagacat cagggcgat ggattgcaga 414
63 attccaagtc ctcatccaa aggtgtcacc aaccttcaga gtcactaagg tccaggctca 474

```

RAW SEQUENCE LISTING

DATE: 03/06/2003

PATENT APPLICATION: US/09/724,000A

TIME: 12:42:46

Input Set : A:\00450Aseq.txt

Output Set: N:\CRF4\03062003\I724000A.raw

```

65 gccacaagt caccatggct cctccagagt aaaagtccaa gattccacct gtgggagcta 534
67 cagatccaga gactttcaag ctgactagag tgcagagaag caagacctca gtgtgatcag 594
69 ccgagactac agcatcttgg gaaccctcag tcagcccca acccctaaca ctttaaccact 654
71 ggtctccaaa ccaacacctg taacttccta atgaaatcat caggaggata ccaaaagaaa 714
73 taaaccataa atcagcatac aactataaaa                                     744

```

76 <210> SEQ ID NO: 2

77 <211> LENGTH: 78

78 <212> TYPE: PRT

79 <213> ORGANISM: Mus musculus

81 <400> SEQUENCE: 2

82 Met Arg Leu Leu Ala Leu Ser Gly Leu Leu Cys Met Leu Leu Leu Cys

83 1 5 10 15

85 Phe Cys Ile Phe Ser Ser Glu Gly Arg Arg His Pro Ala Lys Ser Leu

86 20 25 30

88 Lys Leu Arg Arg Cys Cys His Leu Ser Pro Arg Ser Lys Leu Thr Thr

89 35 40 45

91 Trp Lys Gly Asn His Thr Arg Pro Cys Arg Leu Cys Arg Asn Lys Leu

92 50 55 60

94 Pro Val Lys Ser Trp Val Val Pro Gly Ala Leu Pro Gln Ile

95 65 70 75

98 <210> SEQ ID NO: 3

99 <211> LENGTH: 54

100 <212> TYPE: PRT

101 <213> ORGANISM: Mus musculus

103 <400> SEQUENCE: 3

104 Arg Arg His Pro Ala Lys Ser Leu Lys Leu Arg Arg Cys Cys His Leu

105 1 5 10 15

107 Ser Pro Arg Ser Lys Leu Thr Thr Trp Lys Gly Asn His Thr Arg Pro

108 20 25 30

110 Cys Arg Leu Cys Arg Asn Lys Leu Pro Val Lys Ser Trp Val Val Pro

111 35 40 45

113 Gly Ala Leu Pro Gln Ile

114 50

117 <210> SEQ ID NO: 4

118 <211> LENGTH: 806

119 <212> TYPE: DNA

120 <213> ORGANISM: Homo sapiens

122 <220> FEATURE:

123 <221> NAME/KEY: CDS

124 <222> LOCATION: (29)..(274)

126 <220> FEATURE:

127 <221> NAME/KEY: sig_peptide

128 <222> LOCATION: (29)..(100)

130 <400> SEQUENCE: 4

131 ggaacgaggg aaaatctgcc ttctcacc atg agg ctt cta gtc ctt tcc agc 52

132 Met Arg Leu Leu Val Leu Ser Ser

133 1 5

135 ctg ctc tgt atc ctg ctt ctc tgc ttc tcc atc ttc tcc aca gaa ggg 100

136 Leu Leu Cys Ile Leu Leu Leu Cys Phe Ser Ile Phe Ser Thr Glu Gly

RAW SEQUENCE LISTING

DATE: 03/06/2003

PATENT APPLICATION: US/09/724,000A

TIME: 12:42:46

Input Set : A:\00450Aseq.txt

Output Set: N:\CRF4\03062003\I724000A.raw

```

137      10      15      20
139 aag agg cgt cct gcc aag gcc tgg tca ggc agg aga acc agg ctc tgc 148
140 Lys Arg Arg Pro Ala Lys Ala Trp Ser Gly Arg Arg Thr Arg Leu Cys
141 25      30      35      40
143 tgc cac cga gtc cct agc ccc aac tca aca aac ctg aaa gga cat cat 196
144 Cys His Arg Val Pro Ser Pro Asn Ser Thr Asn Leu Lys Gly His His
145      45      50      55
147 gtg agg ctc tgt aaa cca tgc aag ctt gag cca gag ccc cgc ctt tgg 244
148 Val Arg Leu Cys Lys Pro Cys Lys Leu Glu Pro Glu Pro Arg Leu Trp
149      60      65      70
151 gtg gtg cct ggg gca ctc cca cag gtg tag cactcccaaa gcaagactcc 294
152 Val Val Pro Gly Ala Leu Pro Gln Val
153      75      80
155 agacagcgga gaacctcatg cctggcacct gaggtaccca gcagcctcct gtctccctt 354
157 tcagccttca cagcagttag ctgcaatgtt ggagggttct atctcgggct gcaaggaccc 414
159 tgggaaagt ccagaactcc acgtccttgt ctcaattgtg ccatcaactt tcagagctat 474
161 catgagccaa cctcacccca cagggcctca gtcgccacca tgtgggcctc tccagtgc 534
163 accaccgagc attccaccat gaccggtcac agctacaaat ccagagacca tcaatcctgc 594
165 tagagtgcag ggtggcaagc acccaagggg ggctgaccaa gactgcagag tctcctccat 654
167 cttcagggtc attcagcctc ctggcattta actaccagca tccagtgggc cccaaggaat 714
169 cccttcctag cctcctgaca tgagtctgct ggaaagagca tccaaacaaa caagtaataa 774
171 ataaataaat aaactcaatg cagacacaaa aa 806
174 <210> SEQ ID NO: 5
175 <211> LENGTH: 81
176 <212> TYPE: PRT
177 <213> ORGANISM: Homo sapiens
179 <400> SEQUENCE: 5
180 Met Arg Leu Leu Val Leu Ser Ser Leu Leu Cys Ile Leu Leu Leu Cys
181 1      5      10      15
183 Phe Ser Ile Phe Ser Thr Glu Gly Lys Arg Arg Pro Ala Lys Ala Trp
184      20      25      30
186 Ser Gly Arg Arg Thr Arg Leu Cys Cys His Arg Val Pro Ser Pro Asn
187      35      40      45
189 Ser Thr Asn Leu Lys Gly His His Val Arg Leu Cys Lys Pro Cys Lys
190      50      55      60
192 Leu Glu Pro Glu Pro Arg Leu Trp Val Val Pro Gly Ala Leu Pro Gln
193 65      70      75      80
195 Val
198 <210> SEQ ID NO: 6
199 <211> LENGTH: 57
200 <212> TYPE: PRT
201 <213> ORGANISM: Homo sapiens
203 <400> SEQUENCE: 6
204 Lys Arg Arg Pro Ala Lys Ala Trp Ser Gly Arg Arg Thr Arg Leu Cys
205 1      5      10      15
207 Cys His Arg Val Pro Ser Pro Asn Ser Thr Asn Leu Lys Gly His His
208      20      25      30
210 Val Arg Leu Cys Lys Pro Cys Lys Leu Glu Pro Glu Pro Arg Leu Trp
211      35      40      45

```

RAW SEQUENCE LISTING

DATE: 03/06/2003

PATENT APPLICATION: US/09/724,000A

TIME: 12:42:46

Input Set : A:\00450Aseq.txt

Output Set: N:\CRF4\03062003\I724000A.raw

213 Val Val Pro Gly Ala Leu Pro Gln Val

214 50 55

217 <210> SEQ ID NO: 7

218 <211> LENGTH: 77

219 <212> TYPE: PRT

220 <213> ORGANISM: Rattus norvegicus

222 <400> SEQUENCE: 7

223 Met Arg Leu Leu Thr Leu Ser Gly Leu Phe Phe Met Leu Phe Leu Cys

224 1 5 10 15

226 Leu Cys Val Leu Ser Ser Glu Gly Arg Lys Arg Pro Ala Lys Phe Pro

227 20 25 30

229 Lys Leu Arg Pro Arg Cys His Leu Ser Pro Arg Ser Lys Pro Ile Thr

230 35 40 45

232 Trp Lys Gly Asn His Thr Arg Pro Cys Arg Pro Cys Arg Lys Leu Glu

233 50 55 60

235 Ser Asn Ser Trp Val Val Pro Gly Ala Leu Pro Gln Ile

236 65 70 75

239 <210> SEQ ID NO: 8

240 <211> LENGTH: 4159

241 <212> TYPE: DNA

242 <213> ORGANISM: Homo sapiens

244 <220> FEATURE:

245 <221> NAME/KEY: unsure

246 <222> LOCATION: (160)..(169)

248 <220> FEATURE:

249 <221> NAME/KEY: unsure

250 <222> LOCATION: (3884)..(3893)

252 <220> FEATURE:

253 <221> NAME/KEY: exon

254 <222> LOCATION: (1)..(69)

256 <220> FEATURE:

257 <221> NAME/KEY: exon

258 <222> LOCATION: (2627)..(2725)

260 <220> FEATURE:

261 <221> NAME/KEY: exon

262 <222> LOCATION: (4079)..(4159)

264 <400> SEQUENCE: 8

265 atg agg ctt cta gtc ctt tcc agc ctg ctc tgt atc ctg ctt ctc tgc 48

267 ttc tcc atc ttc tcc aca gaa ggtaggggcag cccccagggg gcagatccct 99

W--> 269 gagcaggatt tcagcatctg ggaagactct gatcaggatt tgttggaggg caggccttgg 159

271 nnnnnnnnnn cgcgcggtact tccagccccc tgggtgaagac gaaagagggc tctttctcct 219

273 gaacctatag gtttggggct caggactgcc tgcagggtggc ttggggggttc cattcacagc 279

275 ccctgcaccc ccaaatacat acccagccta agtaaagtgg tgtgttcgcc atgcaaacac 339

277 acatacaacc tctcagctag attactgtgc ttaagtccta cctatctaga atttctggag 399

279 ccattctctt gtacttgtgt catgcttggg acagagtaaa ttagtggttg gcaaatgaat 459

281 acattaatta gtagaccatc taagtctgaa catcccaaaa cctcatgccc agaaaatata 519

283 catgagcagc tgaaatgaag gtgtgtgtgg tagggagggtg ggggtatgtt atgcatgtt 579

285 agaaggggac accatctttt tacctctata gatatgaata tttagctctc ttgccctttt 639

287 ttcttttttc tttttttttt ttttttttgg atggagtctt gctctgtcac ccaggctgga 699

RAW SEQUENCE LISTING

DATE: 03/06/2003

PATENT APPLICATION: US/09/724,000A

TIME: 12:42:46

Input Set : A:\00450Aseq.txt

Output Set: N:\CRF4\03062003\I724000A.raw

```

289 gtgcagtggc gctatctcag ctactgcaa tctccgcctc ctgggttcaa gcaattctct 759
291 gcctcagcct cccaagtagc tgagattaca ggtgccacc accaagccca gctaattttt 819
293 gtatttttag tacagacagg tttcaccatc ttggccaggc tggctctgaa ctccaaacct 879
295 cgtaatctc ccacctcggc ctcccaaagt gctgggatta caggcgtgag ccaccatgcc 939
297 tggctgcctt tcttgattca gatagctgag tgtttcaatc cttttttctc ttgtctaacc 999
299 ctctagaaac tgcctacatt tttttttgt tttagtgggt atggttactc aaacttttgg 1059
301 gtggggggag ctggagctat agaaatata aaagagaaga aaaacactca attccatgat 1119
303 tcaagagtag ccatgttcaa ctttttggtt atttccttgc atgtagaatt tttaaaaaatt 1179
305 aattgatgta cctatatgtt caaggttata tcttttttat ttatcactat atatatgttt 1239
307 ataatcacc ccaatgctta tgattgaaga tattctggaa gcattttacaa cccagtgtca 1299
309 gcagcagcca tctctgagta gtgggattat aacaagtgtt tgttttacaa agtttctgctg 1359
311 atgaaaatgt cccacatata taataaggaa aacagtgtt agaatctctc ataaacacag 1419
313 cccgtgacat gcaatttatc agacctctat ttttgacat gttggagggt gccagtata 1479
315 ccctagtac aattaaatga ggatagatac cttcccccac aaagtcttct atccatttag 1539
317 gactatctgt agcaaaactc tgaagtagca ttaatcaact aatattttca ggtataactt 1599
319 gctacaagt aacgtactat gatgaattta catgcttaga catttagata gttcacaatt 1659
321 gtgtgctttt ctttttttga agcaagatct tgctctcttg cccaggctcg agtgagtg 1719
323 catgaccacg gctcagtga ggcttgactt ccagggtcga agcaatactc gcacctcagg 1779
325 ttttccagta gctgggaaaa cagggtgcga ccacaatgcc ctgctaattt ttaaaatttt 1839
327 ttgcagagac gaggtctctc taagtgtgcc aggtgtgtct tgaacttctg gactcaagcc 1899
329 atcctccac cttggcctcc cagagtgtca ggatcacagg catgagccac cacacctggc 1959
331 ctactttgca ctttttaatt atgtgtgtaa aggtatatat gtacataaag tatgtccttt 2019
333 attcaggctt tttttctttt tttctttttt ttattttttt gagacgaagt ttttgctctt 2079
335 gttgtccagg ctggagtga atggcatgct cttggctcac cacaacctcc gcctcccggg 2139
337 ttcaagtgat tctcctgcct caacctcctg agtagctggg attacaggca tgcaccaaca 2199
339 tgccaggctg attttgatt tttagttag atggggttcc tccatgttgg tcaggcttgt 2259
341 ctcgaaact cgacctcaag tgatccgccc acctcagcct cccaaagagc taggattaca 2319
343 ggcatgagcc accacaccca gctcagggtt tttttctta ggctagattg ccaaggggag 2379
345 aattattatg tcaaaagaaac tacttattgg acaggaatct gaaaaggatg tgttttgggg 2439
347 ccatgtgtct cccaacattg ttatttctga aaagtaaata acaacaaggc ccaactcttc 2499
349 cctaggacct ctgtagcct ggctcatcct gagtttctct ggataaatat tcctgagccc 2559
351 tgtgccttgg aaggggaagc tcaactcacag acaagcccac taaagacagt ctctcttct 2619
353 ttgtgtc cac cct cag gga aga ggc gtc ctg cca agg cct ggt cag gca 2668
355 gga gaa cca ggc tct gct gcc acc gag tcc cta gcc cca act caa caa 2716
357 acc tga aag gtaagtaccc ccacctcgtc cagactgtgg ggcagaagtt 2765
359 ctacagtggc catgggacca gccacacaca ctgatcagcc cccacccatg gctggcatca 2825
361 ggctctggct gggaggacat ctttgttttg ttgattaatt tgttgactcc ccccaaaaag 2885
363 tcaacaaatt aatcatttta aactgaatac attctgccat ggaaaaaaag caggatgcaa 2945
365 ttagcagatg ttgtgtggaa acacacttac tttaggtgga aggtgtctga gcaggtgaca 3005
367 tttatgagac ctggctcatt tatgagccag gagcctggct gaggcctgtg gaggtggggc 3065
369 atgcaggcag aggaggcagc aagggtgaag ggcaagagtg gggatggaa gacagatggt 3125
371 agcagggctt gagaggtact ccagaaagct aaggaccaa gctgcctgtg aacctgtgg 3185
373 acctggggca cagatcagca tgcaggtcac cagcagggga gtgggcctga ggggtccag 3245
375 agccatagct tggcaggaga taaggcagcc ccagagatgc cagcaggcag catccaggct 3305
377 gcatgaccag aacgaggccc agaagagcaa ggctgccctc tccctgaggc ctggggacac 3365
379 tgggaggcct gtggcggaca ggcccaagct caggagggtc gcgggcaccc agttccctgc 3425
381 acaggggctg caggcccaga gcagatatc actggagttg cccagcccag gtggaagggt 3485
383 caggctgtg gagcttgggt agggcaggca gatcccaaag gggagactgt ggacctgag 3545
385 tcagacagcc tgacaccaac ctggggctcc tgctgaact ctgcagcccc agtgcacct 3605

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/724,000A

DATE: 03/06/2003
 TIME: 12:42:47

Input Set : A:\00450Aseq.txt
 Output Set: N:\CRF4\03062003\I724000A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:8; N Pos. 160,161,162,163,164,165,166,167,168,169,3884,3885,3886,3887
 Seq#:8; N Pos. 3888,3889,3890,3891,3892,3893
 Seq#:22; Xaa Pos. 5,8,10,11,12,14,17,18,19,20,22,25,26,27,31,32,33,34,36,37
 Seq#:22; Xaa Pos. 38,39,40,43,44,46,47,48,49,50,51,52,55,57,59,61,62,64,65
 Seq#:22; Xaa Pos. 66,67,68,69,70,71,81

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23
 Seq#:1; Line(s) 24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43
 Seq#:1; Line(s) 44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63
 Seq#:1; Line(s) 64,65,66,67,68,69,70,71,72,73,74,75,76
 Seq#:2; Line(s) 77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96
 Seq#:2; Line(s) 97,98
 Seq#:3; Line(s) 99,100,101,102,103,104,105,106,107,108,109,110,111,112,113
 Seq#:3; Line(s) 114,115,116,117
 Seq#:4; Line(s) 118,119,120,121,122,123,124,125,126,127,128,129,130,131,132
 Seq#:4; Line(s) 133,134,135,136,137,138,139,140,141,142,143,144,145,146,147
 Seq#:4; Line(s) 148,149,150,151,152,153,154,155,156,157,158,159,160,161,162
 Seq#:4; Line(s) 163,164,165,166,167,168,169,170,171,172,173,174
 Seq#:5; Line(s) 175,176,177,178,179,180,181,182,183,184,185,186,187,188,189
 Seq#:5; Line(s) 190,191,192,193,194,195,196,197,198
 Seq#:6; Line(s) 199,200,201,202,203,204,205,206,207,208,209,210,211,212,213
 Seq#:6; Line(s) 214,215,216,217
 Seq#:7; Line(s) 218,219,220,221,222,223,224,225,226,227,228,229,230,231,232
 Seq#:7; Line(s) 233,234,235,236,237,238,239
 Seq#:8; Line(s) 240,241,242,243,244,245,246,247,248,249,250,251,252,253,254
 Seq#:8; Line(s) 255,256,257,258,259,260,261,262,263,264,265,266,267,268,269
 Seq#:8; Line(s) 270,271,272,273,274,275,276,277,278,279,280,281,282,283,284
 Seq#:8; Line(s) 285,286,287,288,289,290,291,292,293,294,295,296,297,298,299
 Seq#:8; Line(s) 300,301,302,303,304,305,306,307,308,309,310,311,312,313,314
 Seq#:8; Line(s) 315,316,317,318,319,320,321,322,323,324,325,326,327,328,329
 Seq#:8; Line(s) 330,331,332,333,334,335,336,337,338,339,340,341,342,343,344
 Seq#:8; Line(s) 345,346,347,348,349,350,351,352,353,354,355,356,357,358,359
 Seq#:8; Line(s) 360,361,362,363,364,365,366,367,368,369,370,371,372,373,374
 Seq#:8; Line(s) 375,376,377,378,379,380,381,382,383,384,385,386,387,388,389
 Seq#:8; Line(s) 390,391,392,393,394,395,396,397,398,399,400,401,402,403,404
 Seq#:8; Line(s) 405,406,407,408
 Seq#:9; Line(s) 409,410,411,412,413,414,415,416,417,418,419,420,421
 Seq#:10; Line(s) 422,423,424,425,426,427,428,429,430,431,432,433,434
 Seq#:11; Line(s) 435,436,437,438,439,440,441,442,443,444,445,446,447
 Seq#:12; Line(s) 448,449,450,451,452,453,454,455,456,457

VERIFICATION SUMMARY

DATE: 03/06/2003

PATENT APPLICATION: US/09/724,000A

TIME: 12:42:47

Input Set : A:\00450Aseq.txt

Output Set: N:\CRF4\03062003\I724000A.raw

L:271 M:258 W: Mandatory Feature missing, <223> Tag not found for SEQ ID#:8

L:271 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:159

M:341 Repeated in SeqNo=8

L:816 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22 after pos.:0

M:341 Repeated in SeqNo=22